



A PLANAR VERTICAL CHANNEL DMOS STRUCTURE

Richard A. Blanchard

2

3

4 FIELD OF THE INVENTION

5 This invention relates to a double diffused MOS (DMOS)
6 transistor having a vertical channel region, and in
7 particular to a planar DMOS transistor having a vertical
8 gate.

9

10 BACKGROUND OF THE APPLICATION

11 Double diffused MOS (DMOS) transistors are well known
12 in the prior art. For example, U.S. Patent No. 4,344,081,
13 issued to Pao et al. on August 10, 1982, which is
14 incorporated herein by reference, shows one such prior art
15 structure. Fig. 1 shows a cross section of a prior art
16 N-channel DMOS power transistor. This prior art structure
17 includes an N₃₀ epitaxial layer 11 formed on an N₃₀⁺ silicon
18 substrate 10. Gate oxide layer 16 is formed on epitaxial
19 layer 11 and doped polysilicon gate 15 is formed on oxide
20 layer 16. Oxide layer 9 covers gate 15. P-type body
21 regions 12a and 12b are diffused into epitaxial layer 11,
22 and N₃₀⁺ source regions 13a and 13b are diffused into body
23 regions 12a and 12b, respectively. Source regions 13a and
24 13b are electrically tied to body regions 12a and 12b by
25 metal contacts 18 and 19, respectively. Contacts 18 and 19
26 are also electrically tied together. Regions 12c1 and 12c2
27 beneath gate 15 in body regions 12a and 12b, respectively,
28 are channel regions. When the potential between gate 15 and
29 source regions 13a and 13b is sufficiently high and with a
30 positive voltage on drain contact 17, carriers flow
31 laterally from source regions 13a and 13b through channel
32 regions 12a and 12b, respectively, to drain region 11 and
33 then vertically downward through drain region 11 and N₃₀⁺
34 substrate 10 to drain contact 17, as indicated by arrows 20a
35 and 20b in Fig. 1. P-channel DMOS transistors have a
36 similar structure, but P-type and N-type regions are
37 reversed, and a voltage of the opposite sign produces
38 current flow.

1 As explained above, the carriers that flow in the prior
2 art vertical DMOS transistors shown in Fig. 1 must change
3 direction, first flowing laterally and then vertically.
4 Carrier flow is more efficient if the source, body and drain
5 regions are arranged vertically as shown in Fig. 2. Fig. 2
6 shows a cross section of a prior art DMOS transistor with a
7 U shaped gate extending into the epitaxial layer. This
8 structure is due to Ueda et al. and is explained in more
9 detail in A New Vertical Power MOSFET Structure with
10 Extremely Reduced On-Resistance, IEEE TRANSACTIONS ON
11 ELECTRON DEVICES, VOL. ED-32, NO. 1, January 1985, which is
12 incorporated herein by reference. In this prior art
13 structure N_{30}^- epitaxial layer 11 is again formed on N_{30}^+
14 substrate 10. A P-type dopant is diffused into epitaxial
15 layer 11 and an N_{30}^+ -type dopant is diffused into a portion of
16 the epitaxial layer that has been doped with a P-type
17 dopant. Rectangular groove 23, having vertical walls, is
18 then etched in the epitaxial layer using reactive ion beam
19 etching, thereby creating P-type body regions 20a and 20b
20 and corresponding N_{30}^+ source regions 21a and 21b as shown in
21 Fig. 2. Source regions 21a and 21b are electrically tied to
22 body regions 20a and 20b, respectively, by metal contacts 18
23 and 19 which are also electrically tied together. A slight
24 wet etch is then applied to smooth the surface of groove
25 23. Gate oxide 24 is formed in rectangular groove 23, and a
26 U shaped polysilicon gate 25 is formed over gate oxide 24.

27 The prior art structure of Fig. 2 has the advantage
28 that when the gate to source potential is sufficient to turn
29 on the transistor, carriers flow vertically from N_{30}^+ source
30 regions 21a and 21b through channel regions 22c1 and 22c2 in
31 body regions 20a and 20b, respectively, and continue to flow
32 vertically downward through drain region 11 to N_{30}^+ substrate
33 10 and drain contact 17. However, the structure of Fig. 2
34 has a disadvantage in that it is difficult to fabricate
35 because it requires the formation of a U-shaped gate and
36 results in a transistor with a nonplanar surface.

37

38

1 SUMMARY OF THE INVENTION

2 *p* A DMOS device is disclosed which has a vertical gate
3 and a planar surface. The device has a flat surface for all
4 masking steps while still allowing contact to be made to the
5 vertical gate.

6 In one embodiment, a DMOS power transistor is disclosed
7 which has a drain region of a first conductivity type, a
8 body region of a second conductivity type formed above the
9 drain region, and a source region of first conductivity
10 type. An upward opening rectangular groove extends downward
11 through the source and body regions and into the drain
12 region so that a first source region in a first body region
13 lies on one side of the rectangular groove and a second
14 source region in a second body region lies on the other side
15 of the rectangular groove.

16 The upward opening rectangular groove is lined with an
17 upward opening dielectric region which is filled with the
18 gate region so that a vertical gate is formed having a top
19 surface which lies between the first and second source
20 regions. An insulating layer is then formed over the above
21 structure so that a transistor with a planar surface is
22 obtained. In operation, carriers flow vertically between
23 the source and drain regions.

24 In another embodiment, a semiconductor device having a
25 vertical gate region is formed in a block of semiconductor
26 material. The vertical gate region lies in an upward
27 opening dielectric region which lines an upward opening
28 rectangular groove. A source region of a first conductivity
29 type is formed above a body region of a second conductivity
30 type which lies above a drain region of a first conductivity
31 type. The source, body and drain regions are all adjacent
32 one vertical surface of the dielectric material. The top
33 surface of the vertical gate region lies opposite the source
34 region and the bottom surface of the gate region lies
35 opposite the drain region. The second embodiment also
36 includes an insulating layer formed over the gate, source
37 and body regions resulting in a device with a planar top
38 surface. Still other embodiments are described below.

DE BRIEF DESCRIPTION OF THE DRAWINGS

2 p Figure 1 shows a prior art N-channel DMOS transistor;
3 p Figure 2 shows a prior art DMOS transistor having
4 vertical channel regions;
5 p Figure 3 shows one embodiment of the DMOS structure of
6 the present invention;
7 p Figures 4a through 4f show process steps in the
8 formation of the transistor shown in Fig. 3;
9 p Figure 5 shows a second embodiment of the DMOS
10 transistor of the present invention;
11 p Figure 6 shows a third embodiment of the DMOS
12 transistor of the present invention;
13 p Figure 7 shows a cross-section of an insulated gate
14 transistor formed according to the present invention;
15 p Figure 8 shows a cross-section of an MOS-gated silicon
16 controlled rectifier formed according to the present
17 invention; and
18 p Figure 9 shows a top view of one surface geometry
19 employed by the present invention.

20

DE DETAILED DESCRIPTION OF THE INVENTION

22 p Figure 3 shows one embodiment of the vertical gate
23 planar DMOS power transistor of the present invention. The
24 process sequence for fabricating this N-channel transistor
25 is shown in Figs. 4a-4f. In other embodiments, the vertical
26 gate planar DMOS transistor of the present invention is a
27 P-channel device.

28 Substrate 10 shown in Fig. 4a is a silicon wafer doped
29 with N-type impurities so that its resistivity is within the
30 range of 0.005 to 0.1 ohm-cm (in one embodiment, 0.02 ohm-cm
31 cm). An N-type epitaxial layer 11 having a resistivity
32 between 0.2 and 100 ohm-cm (in one embodiment, 2.4 ohm-cm)
33 is then grown on substrate 10 to a thickness between 6 and
34 150 microns (in one embodiment, 13.5 μ). Substrate 10 and
35 epitaxial layer 11 typically have a [100] crystal
36 orientation.

37 A layer of dielectric material (not shown) is then
38 formed over the wafer by heating the wafer in an oxygen

1 atmosphere at about $900\text{--}1200^\circ\text{C}$ to form a silicon dioxide
2 layer approximately $1,000\text{--}10,000\text{\AA}$ thick over the surface of
3 the wafer. Using standard photoresist techniques, a body
4 mask pattern is transferred to the surface of the silicon
5 dioxide layer which exposes those regions of the silicon
6 dioxide layer through which ions are implanted to form P
7 region 20. (In one version, no body mask is needed, and
8 P-type dopant is implanted to form region 20.) In one
9 embodiment, P region 20 is formed by implanting boron ions
10 at a dosage between 10^{13} and 2×10^{14} ions/cm² at an energy
11 level between 40 and 120 KEV and then annealing the
12 structure for approximately 4 to 12 hours in an atmosphere
13 of oxygen or nitrogen. P region 20 typically ranges in
14 depth from $\frac{1}{2}$ to $\frac{4}{3}$ microns. Alternatively, P region 20 may
15 be formed using standard diffusion techniques.

16 After a source mask is applied to the wafer, the wafer
17 may then be etched to thin or remove the silicon dioxide
18 formed during the annealing process. (If the oxide formed
19 after the body diffusion is not too thick, i.e., it does not
20 block the implant, no thinning etch is needed.) An N-type
21 ion implantation is performed using arsenic or phosphorus
22 ions at a dosage between 5×10^{14} and 1×10^{16} ions/cm²
23 using an implant energy of 50 to 150 KEV. The wafer is then
24 annealed at a temperature of $900\text{--}1200^\circ\text{C}$ for approximately
25 0.5 to 3 hours in an atmosphere of oxygen and nitrogen to
26 form N₂⁺ source region 21 which typically ranges in depth
27 from $\frac{1}{2}$ to 2 microns.

28 The silicon dioxide layer 30 is formed during the above
29 source drive-in. A gate groove mask (not shown) is applied,
30 and the SiO₂ layer is etched using buffered HF.

31 As shown in Fig. 4b, rectangular groove 31 having
32 vertical sidewalls is then etched using reactive ion beam-
33 etching or other etching technique which permits anisotropic
34 etching regardless of crystallographic orientation. Groove
35 31 preferably has a width less than or equal to $2\mu\text{m}$ and a
36 depth between $3\mu\text{m}$ and $10\mu\text{m}$. As shown in Fig. 4c, gate
37 dielectric layer 32 is then formed over the surfaces of the
38 groove 31. In one embodiment, gate dielectric layer 32 is

1 silicon dioxide having a thickness in a range of $500\text{--}1000\text{\AA}$
 2 and is formed by heating the wafer in an oxygen ambient ⁵²¹
 3 containing water at a temperature of $900\text{--}1100^\circ\text{C}$ for 0.5 to
 4 two hours (in another embodiment, insulating layer 32 is a
 5 combination of silicon dioxide and silicon nitride which is
 6 either grown or deposited). The gate dielectric forms an
 7 inner, upward opening, rectangular groove 31*.
 8 Polysilicon layer 33 (shown in Fig. 4d) is then
 9 deposited using a low pressure chemical vapor deposition
 10 process (LPCVD) to a thickness sufficient to fill
 11 rectangular groove 31*. For example, if groove 31 is 1.5
 12 microns wide and 6 microns deep, polycrystalline silicon
 13 layer 33 is deposited with a thickness of $1.2\text{ }\mu\text{m}$.
 14 Polycrystalline silicon layer 33 is doped ⁸²either during
 15 deposition or subsequent to deposition, typically using
 16 phosphorus, so that it has a sheet resistance of between 30
 17 and 50 ohms/square . Alternatively, in another embodiment,
 18 layer 33 comprises a layer of silicide formed using
 19 conventional techniques to a depth sufficient to fill
 20 rectangular groove 31*.
 21 Polycrystalline silicon layer 33 is then subjected to a
 22 CF_4 etch or another etch technique without using a mask in
 23 the trenched area, groove 31*, (except for a mask (not
 24 shown) which may be placed at any convenient point along the
 25 length of groove 31 in order to keep a contact pad (not
 26 shown) to the to-be-formed gate 34 shown in Fig. 4e) in
 27 order to remove the polycrystalline silicon not in groove
 28 31*. The portion of polycrystalline silicon layer 33
 29 remaining in groove 31* after the CF_4 etch is denoted by 34
 30 in Fig. 4e and serves as the gate of the vertical DMOS
 31 transistor. The etch is continued until top surface 34a
 32 lies $0.25\text{--}0.5\text{ }\mu\text{m}$ below the top surface of layer 30. This top
 33 surface depth is controlled by etch time past clearing the
 34 field. The etch must be terminated so that gate 34 overlaps
 35 N_2^+ regions 21a and 21b shown in Fig. 4f after the subsequent
 36 oxidizing step. The wafer is then oxidized in an atmosphere
 37 containing oxygen (which consumes a portion of polysilicon
 38 layer 33 in groove 31*) until the top surface of the

1 oxidized portion 35 above gate 34 forms an essentially flat
2 (planar) surface with the top surface of passivating layer
3 30 whose thickness may also be slightly increased during the
4 formation of region 35. Of importance, the etch to form
5 surface 34a must be terminated sufficiently soon so that
6 after the oxidation which forms silicon dioxide layer 35,
7 the top portion of gate 34 overlaps N_{30}^+ source regions 21a
8 and 21b (see Fig. 4f).

9 The above structure has a flat surface for all masking
10 steps while still allowing contact to be made to the gate
11 region. The source/body-contact shown schematically in Fig.
12 3 is fabricated using prior art techniques, and in cross
13 section typically appears as shown in Fig. 1.

14 When the gate-to-source potential is sufficiently high
15 and with a positive potential on drain 17 (Fig. 3),
16 electrons flow vertically from N_{30}^+ source regions 21a and 21b
17 through channel regions 22c1 and 22c2 in body regions 20a
18 and 20b, respectively, and continue to flow vertically
19 downward through drain 11 and N_{30}^+ substrate (drain) 10 to
20 drain contact 17.

21 Typically, many DMOS devices similar to the one shown
22 in cross section in Fig. 3 are formed simultaneously.
23 Layout efficiency varies with surface geometry. There is a
24 wide variety of layouts. Fig. 9 shows a top view of one
25 surface geometry employed by this invention, namely, a
26 square source and body region on a square gate grid 35. In
27 Fig. 9, S denotes the locations of the source regions, B the
28 locations of the body regions, and G the locations of the
29 gate regions. The dotted line shown in Fig. 9 corresponds
30 to the cross section shown in Fig. 3. In another layout,
31 (not shown) the gate and source and body regions are
32 interdigitated. Another layout (not shown) has hexagonal
33 source and body regions on an hexagonal gate grid. Still
34 another layout employs square source and body regions on a
35 hexagonal gate grid. The latter layout is more efficient
36 than the others. Other source geometries include
37 rectangles, circles and triangles.

38 The structure shown in Fig. 3 reduces the total area

1 requirement from 30% to 50% below that of the Ueda device
2 shown in Fig. 2.

3 Fig. 5 shows a cross section of an alternate embodiment
4 of the invention in which a P^- epitaxial layer 40 is formed
5 on substrate 10 in place of N^- epitaxial layer 11 and in
6 which gate 34 in groove 31 reaches through to the N^+
7 substrate. In this embodiment the P^- epitaxial layer serves
8 as the body region of the transistor. This embodiment
9 results in a transistor having a lower on resistance than
10 the device shown in Fig. 4a by as much as a factor of 2 and
11 a lower breakdown voltage (typically around 30 volts) than
12 the device shown in Fig. 4a.

13 Fig. 6 shows another alternate embodiment of the
14 invention which is similar to the embodiment shown in Fig.
15 4f except that groove 31 is formed sufficiently deep so that
16 gate 34 extends completely through the epitaxial layer 11
17 and into substrate region 10. This alternate embodiment
18 also has a lower on resistance and a lower breakdown voltage
19 (typically about 30 volts) than the device shown in Fig.
20 4a. This is acceptable in low voltage applications, for
21 example, low voltage motors and Schottky diode replacements.

22 The above description has been given in terms of DMOS
23 transistors, but the invention also applies to other MOS-
24 gated devices such as an MOS-gated SCR or a MOS-gated
25 conductivity modulated device.

26 Fig. 7 shows a cross section of one embodiment for an
27 insulated gate transistor. Fig. 8 shows a cross section of
28 one embodiment for an MOS-gated silicon controlled
29 rectifier.

30 The silicon controlled rectifier shown in Fig. 8 is
31 fabricated in the same manner as explained above for the
32 transistor shown in Fig. 3 in connection with Figs. 4a
33 through 4f except that the starting material is a silicon
34 substrate 41 heavily doped with P-type material, for example
35 doped with Boron to a resistivity of 0.01 ohm-cm.

36 In operation, silicon controlled rectifier 60 is
37 switched on by appropriately biasing source/body terminal 50
38 (shown schematically in Fig. 8), which provides electrical

1 contact to source regions 21a and 21b and body regions 20a
2 and 20b, gate terminal 49, which contacts gate 34, and
3 substrate contact 51. Gate contact 49 and substrate contact
4 51 are biased positive relative to source/body contact 50.
5 MOSFET mode conduction is initiated by electrons flowing
6 from source regions 21a and 21b through channel regions 22q1
7 and 22q2, respectively, to N_{31}^- drain region 11.

8 Drain region 11 also serves as the base for the PNP
9 bipolar (junction) transistors comprising emitter region
10 20a, base 11, and collector region 41; and emitter region
11 20b, base 11, and collector region 41, respectively. With
12 substrate contact 51 biased positively with respect to
13 source/body contact 50 and with electrons flowing into base
14 11, the bipolar transistors are triggered into conduction
15 and may latch in the on state even when the gate bias is
16 removed.

17 The insulated gate transistor shown in Fig. 7 is
18 similar to the structure shown in Fig. 8 except that in
19 order to suppress thyristor action, the IGT shown in Fig. 7
20 is designed with narrow N_{30}^+ source regions 21a and 21b which
21 reduce the lateral body resistance beneath the source
22 regions. See The Insulated Gate Transistor: A New
23 Three-Terminal MOS-Controlled Bipolar Power Device, IEEE
24 Trans. on Electron Devices, Vol. ED-31 No. 6, June 1984,
25 which is incorporated herein by reference. For example, the
26 width w_{40}^+ of N_{30}^+ source regions 21a and 21b in Fig. 7 is
27 typically between 2 μ m and 4 μ m whereas the width w of N_{30}^+
28 source regions 21a and 21b in Fig. 8 is typically between
29 6 μ m and 8 μ m. The doping profile of the body regions beneath
30 source regions 21a and 21b in Fig. 7 is also selected to
31 reduce lateral body resistance beneath source regions 21a
32 and 21b. See Blanchard, U.S. Patent No. 4,345,265, issued
33 August 17, 1982, which is incorporated herein by
34 reference. A low lateral body resistance beneath source
35 regions 21a and 21b in Fig. 7 prevents the NPN transistor
36 formed by source regions 21a and 21b, body regions 22c1 and
37 22c2, and the N-type drain region 11 from becoming active.
38 As long as this NPN bipolar transistor does not turn on, the

1 regenerative action characteristic of an SCR does not occur.
2 The above embodiments are meant to be exemplary and not
3 limiting. In view of the above disclosure, many
4 modifications and substitutions will be obvious to one of
5 average skill in the art without departing from the scope of
6 the invention.

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

am I doing;